

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

----- x
ELSA GULINO, MAYLING RALPH, PETER WILDS, and :
NIA GREENE, on behalf of themselves and all others :
similarly situated, : 96 Civ. 8414 (KMW)
:
Plaintiffs, :
:
- against - :
:
THE BOARD OF EDUCATION OF THE CITY SCHOOL :
DISTRICT OF THE CITY OF NEW YORK; and NEW :
YORK STATE EDUCATION DEPARTMENT, :
:
Defendant. x

**PLAINTIFFS' REMAND SUBMISSION
PURSUANT TO THIS COURT'S DECEMBER 8, 2009 ORDER**

Anjana Samant
CENTER FOR CONSTITUTIONAL RIGHTS
666 Broadway, Seventh Floor
New York, New York 10012
(212) 614-6445 (Telephone)
(212) 614-6499 (Facsimile)

Joshua S. Sohn
Anthony D. Gill
Spencer Stiefel
Rachel V. Stevens
DLA PIPER LLP (US)
1251 Avenue of the Americas
New York, New York 10020
(212) 335-4500 (Telephone)
(212) 335-4501 (Facsimile)

Stephen G. Seliger
Joel Hellman
155 North Michigan Avenue, Suite 501
Chicago, Illinois 60601
(312) 616-4244 (Telephone)
(312) 427-1850 (Facsimile)

Attorneys for Plaintiffs

TABLE OF CONTENTS

	PAGE
PRELIMINARY STATEMENT	1
FACTUAL AND PROCEDURAL HISTORY	5
A. Background	5
B. Standard Of Job Relatedness.....	7
C. Pervasive Lack Of Documentation And Well-Established Record	8
ARGUMENT	10
I. JUDGE MOTLEY CORRECTLY CONCLUDED THAT THE LAST DOES NOT MEET THE REQUIREMENTS OF GUARDIANS	10
A. NES Did Not Perform A Suitable Job Analysis.....	11
B. NES Did Not Exercise Reasonable Competence In Test Construction	15
1. The Composition And Work Of The Content Advisory Committee And The Bias Review Committee Were Insufficient	15
2. The Work Of The Equity Advisory Board Was Insufficient.....	18
3. Item Drafting For The LAST Did Not Meet Professionally Acceptable Standards.....	19
4. Pilot Testing Of The LAST Did Not Meet Professionally Acceptable Standards.....	20
5. Defendant Violated Professional Standards By Failing To Conduct Continuing Procedures To Eliminate Bias.....	22
6. The Defendant Presented No Evidence Of Purported Changes To The LAST Based On New Learning Standards In The Late 1990s.....	24
7. Defendant's Past Justification For The Poor Test Construction Of The LAST – That Students Of Certified Teachers Perform Better In The Classroom – Has No Basis In The Record	24
C. The Content Of The LAST Is Not Related To The Content Of The Job	25
D. The Content Of The LAST Is Not Representative Of The Content Of The Job	28
E. The Scoring System Of The LAST Does Not Distinguish Between Those Who Will And Will Not Cause Harm In The Classroom	30
II. THE LAST AND CORE BATTERY WERE MISUSED TO MAKE EMPLOYMENT DECISIONS REGARDING IN-SERVICE TEACHERS	33
A. The LAST	33
B. The Core Battery	35

TABLE OF CONTENTS
(continued)

	PAGE
1. The Second Circuit Did Not Consider Whether The Core Battery Was Misused And Therefore The Issue Remains Open On Remand.....	35
2. The Core Battery Was Misused	36
III. NO FURTHER EVIDENTIARY HEARINGS WOULD BE USEFUL OR PROPER	37
CONCLUSION.....	38

TABLE OF AUTHORITIES

CASES	PAGE
<i>Albemarle Paper Co. v. Moody</i> , 422 U.S. 405 (1975).....	7, 15, 32
<i>AMAE v. California</i> , 937 F. Supp. 1397 (N.D. Cal. 1996), <i>aff'd</i> , 231 F.3d 572 (9th Cir. 2000)	14
<i>Berkman v. City of N.Y.</i> , 536 F. Supp. 177 (E.D.N.Y. 1982), <i>aff'd</i> , 705 F.2d 584 (2d Cir. 1983).....	10
<i>Bernard v. Gulf Oil</i> , 841 F.2d 547 (5th Cir. 1988)	34
<i>Burrell v. U.S.</i> , 467 F.3d 160 (2d Cir. 2006).....	35, 36
<i>County of Suffolk v. Stone & Webster Eng'g Corp.</i> , 106 F.3d 1112 (2d Cir. 1997).....	9
<i>Cuesta v. NY State Office of Court Admin.</i> , 657 F. Supp. 1084 (S.D.N.Y. 1987).....	25, 26
<i>Dothard v. Rawlinson</i> , 433 U.S. 321 (1977).....	26
<i>EEOC v. Atlas Paper Box Co.</i> , 868 F.2d 1487 (6th Cir. 1989)	34
<i>Fickling v. N.Y. State Department of Civil Service</i> , 909 F. Supp 185 (S.D.N.Y. 1995).....	23
<i>Georgiadis v. First Boston Corp.</i> , 167 F.R.D. 24 (S.D.N.Y. 1996)	38
<i>Guardians Ass'n v. Civil Service Comm'n of N.Y.</i> , 630 F.2d 79 (2d Cir. 1980).....	passim
<i>Gulino v. N.Y. State Educ. Dep't</i> , 460 F.3d 361 (2d Cir. 2006).....	passim
<i>Haouari v. U.S.</i> , 510 F.3d 350 (2d Cir. 2007).....	38
<i>Lanning v. SEPTA</i> , 181 F.3d 478 (3d Cir. 1999).....	26

TABLE OF AUTHORITIES
(continued)

	PAGE
<i>Meacham v. Knolls Atomic Power Lab.</i> , No. 09-2037-CV, 2009 WL 4912205 (2d Cir. 2009)	35, 36
<i>NAACP v. Town of East Haven</i> , 70 F.3d 219 (2d Cir. 1995).....	10
<i>New York City Trans. Auth. v. Beazer</i> , 440 U.S. 568 (1979).....	7
<i>Rosario v. Ercole</i> , 582 F.Supp. 2d 541 (S.D.N.Y. 2008).....	38
<i>U.S. v. Ben Zvi</i> , 242 F.3d 89 (2d Cir. 2001).....	9, 35
<i>U.S. v. Minicone</i> , 994 F.2d 86 (2d Cir. 1993).....	35, 36
<i>U.S. v. Vulcan Soc., Inc.</i> , 637 F. Supp. 2d 77 (E.D.N.Y. 2009)	28
<i>United States ex rel. Bisordi v. LaVallee</i> , 461 F.2d 1020 (2d Cir. 1972) (holding remand to district court)	38
<i>Vulcan Soc. of New York City Fire Dep't v. Civil Serv. Comm'n</i> , 490 F.2d 387 (2d Cir. 1973).....	26
<i>Walston v. County School Bd. of Nansemond County, Virginia</i> , 492 F.2d 919 (4th Cir. 1974)	33
<i>Walston v. County School Bd. of Nansemond County, Virginia</i> , 566 F.2d 1204 (4th Cir. 1977)	34, 37
<i>Watson v. Fort Worth Bank & Trust</i> , 487 U.S. 977 (1988).....	7
<i>Williams v. Ford Motor Co.</i> , 187 F.3d 533 (6th Cir Mich. 1999)	17
<i>York v. Alabama State Bd. of Ed.</i> , 581 F. Supp. 779 (M.D. Ala. 1983)	34

TABLE OF AUTHORITIES
(continued)

STATUTES	PAGE
42 U.S.C. §§ 2000e-2000e-17 (Title VII of the Civil Rights Act of 1964)	passim
29 C.F.R. 1607.15(A)(3).....	9
29 C.F.R. § 1607.9(A)	11, 15, 19
29 C.F.R. § 1607.16(D)	10
29 C.F.R. § 1607.149 (C)(5).....	32

Plaintiffs, a class of African-American and Latino New York City public school teachers who claim Defendant's use of the Liberal Arts and Sciences Test ("LAST") and NTE Core Battery ("Core Battery") violated Title VII, respectfully submit the following points of law and fact through their attorneys pursuant to this Court's December 8, 2009 Order directing the parties to address: (1) whether Judge Motley was correct that under the standard established by *Guardians Association v. Civil Service Commission of New York* ("*Guardians*")¹ the LAST is not job related under Title VII of the Civil Rights Act of 1964;² (2) whether the LAST was misused for employment decisions regarding in-service teachers; (3) whether the misuse of the Core Battery remains an open issue on remand; (4) whether the Core Battery was misused for employment decisions regarding in-service teachers; and (5) whether the Court can resolve this case without holding evidentiary hearings to determine credibility of witnesses.

PRELIMINARY STATEMENT

The answer to each of the questions above is yes. After five months of trial, testimony from more than forty witnesses, the introduction of more than 1,000 exhibits, and a record spanning 3,600 pages of trial transcript, the Honorable Constance Baker Motley concluded that under the standards established by the Second Circuit in *Guardians*, "plaintiffs would emerge triumphant with respect to the LAST" on their Title VII claims.³ Nothing in the intervening seven years changes Judge Motley's well-founded conclusion: under Title VII and *Guardians* the LAST is not job related and thus Plaintiffs must "emerge triumphant." In addition, the evidence in the record unequivocally establishes that the LAST and Core Battery (collectively, the

¹ 630 F.2d 79 (2d Cir. 1980); see also *Gulino v. New York State Education Department*, 460 F.3d 361, 384-85 (2d Cir. 2006).

² See 42 U.S.C. §§ 2000e-2000e-17.

³ February 17, 2010 Affirmation of Joshua Sohn ("Sohn Aff."), Remand Exh. A (September 4, 2003 Findings of Fact and Conclusions of Law by The Honorable Constance Baker Motley (the "Findings")), ¶ 161.

“Tests”) were each misused—rendering them invalid under Title VII—when they were used as the basis for employment decisions related to in-service teachers. Finally, no additional evidentiary hearings are necessary to determine credibility as all evidence needed to resolve the issues before the Court exists on the face of the current record.

The Tests are certification examinations that were required at different times in order to become a full-time, permanently appointed teacher in New York City’s public schools. From a test-development perspective, these certification tests should have been designed to distinguish between applicants who were minimally competent to do the job of a New York City public school teacher (*i.e.* those who will not cause harm in the classroom), and those applicants who were incompetent (*i.e.* those who will cause harm in the classroom). The Tests, however, were neither developed, nor used in a manner consistent with this purpose.

The test developer for the LAST, National Evaluation Systems (“NES”), proposed a multi-stage approach to develop the LAST that was supposed to include the following steps:

- A “task analysis” of the tasks performed by a teacher, which was to be supported by curricula material and textbooks from teacher colleges and public schools in New York as well as professional literature on education.
- The task analysis was also to be supported by interviews with New York public school teachers, educational experts, as well as deans and faculty administrators at New York colleges.
- Based on the task analysis, NES contemplated developing a framework and objectives for the LAST, which would comprise the core subjects on which teachers would be tested.
- Once the framework and objectives were drafted, NES contemplated a review by three committees: the Equity Advisory Board (“EAB”); the Content Advisory Committee (“CAC”); and the Bias Review Committee (“BRC”). Together, the purpose of these committees was to review the framework and objectives, and later individual test questions, and to develop coherent and bias-free content.
- Once the objectives were developed, a Job Analysis Survey was to be sent to New York college faculty and public school teachers, asking them to weight each objective on a scale of 1 (no importance) to 5 (very great importance).

- After the objectives were confirmed through the Job Analysis Survey, individual test questions would be drafted. These questions were to be field tested and reviewed by the CAC, BRC, and EAB.
- Finally, the questions would be assembled into pilot tests and administered across the state. Based on the results of these pilot tests, a cut score would be established and then the test would be administered publicly.

Although seemingly thorough on its face, the process above was deeply flawed in its execution and the record demonstrates that the actual process cannot satisfy the job-relatedness validation standards of *Guardians*.

Under *Guardians*, the Board of Education of the City School District of the City of New York (“BOE” or Defendant) bears the burden of persuasion and production to establish five elements,⁴ each of which is necessary to validate a challenged test,⁵ and the failure to establish any of which is fatal to BOE’s case. The expansive trial record makes plain, however, that BOE cannot establish any, much less all, of these elements.

- First, no professionally acceptable job analysis was conducted. Insufficient evidence of the background materials allegedly collected or interviews allegedly conducted was presented to demonstrate a reasoned development of the knowledge, skills, and abilities (“KSAs”) possessed by those teachers who were minimally competent to perform the job of public school teacher. (*See Sec. I(A), infra.*)
- Second, reasonable competence was not used in the construction of the LAST. Among the many failings in the construction of the LAST was the failure to document the work of various committees intended to keep bias from the exam. In addition, NES failed to draft test items, conduct pilot tests, or continually monitor the LAST for bias in compliance with professional standards. Such failings preclude a finding of validity. (*See Sec. I(B), infra.*)
- Third, the content of the LAST is not related to the content of the job. Only where a party establishes that test content is “essential to effective job

⁴ See *Gulino*, 460 F.3d at 382 (“Disparate impact claims follow a three-part analysis involving shifting evidentiary burdens . . . Once the plaintiff has established a prima facie case . . . the defendant may rebut a plaintiff’s prima facie showing by ‘demonstrating that the challenged practice is job related for the position in question and consistent with business necessity’ (citation omitted).)

⁵ *Id.* at 384-85 (listing five elements of job relatedness test under *Guardians*).

performance,” is it related to the content of a job. Here, the failure to conduct a suitable job analysis that identified the KSAs required of a minimally competent teacher precludes a finding the LAST is related to the job. Moreover, NES failed to define the domain of the LAST, making a comparison to the content of the job impossible. (*See Sec. I(C), infra.*)

- Fourth, the LAST is not representative of the content of the job. NES purportedly mailed surveys to New York teachers asking them to weight the “importance” of the LAST objectives. Yet, NES did not provide a definition of “importance” to the survey participants. In addition, witnesses confirmed there is no empirical evidence that supports the weighting given to the essay section or other test questions. Thus, BOE cannot establish the content of the test represents proportionally the KSAs of the job. (*See Sec. I(D), infra.*)
- Fifth, and finally, the scoring of the LAST was not tied to any analysis of who would be a minimally competent teacher, *i.e.*, a teacher who would not cause harm. Although both experts that testified at trial agreed that the purpose of a certification exam like the LAST is to prevent a specific harm,⁶ evidence at trial revealed that no one defined the concept of harm in the context of the LAST. Without a definition of harm there was no context in which to determine where a cut-score should be set. Moreover, the process for establishing the cut-score that was used failed to meet professional standards. Because no valid cut score was developed, the test is not legal under Title VII. (*See Sec. I(E), infra.*)

The evidence in the record plainly demonstrates that BOE cannot satisfy the job-relatedness validation standards set by *Guardians*.

In addition, even if the Court were to disregard the overwhelming evidence that the LAST was not properly developed and validated, both the LAST and Core Battery were misused in violation of Title VII as the basis for employment decisions regarding in-service teachers. Professional standards do not permit the use of a test for any purpose other than that for which it was intended and validated. Here, the developers of both the LAST and Core Battery testified at trial that the intended purpose of the examinations was for certification only and not for the purpose of making employment decisions, including decisions concerning salaries, pensions, or seniority. The evidence also established unequivocally that this direction was simply ignored

⁶ Findings, ¶ 147 (citing testimony of Dr. Frank Landy and Dr. William Mehrens).

and the Tests were used for precisely those prohibited purposes. Because the LAST and Core Battery were misused by BOE, Plaintiffs are entitled to judgment under Title VII. (*See Sec. II, infra.*)

Finally, no additional evidence or witness testimony is necessary to resolve the issues before this Court. Even crediting BOE and the New York State Education Department's ("SED") witnesses' trial testimony, the evidence establishes both the failure to validate the LAST, as well as the misuse of both the LAST and Core Battery. Moreover, sudden "supplemental" recollections by these witnesses more than seven years after trial that would so dramatically change the existing record would be inherently incredible. (*See Sec. III, infra.*) The Court should find Defendant liable under Title VII and proceed to the next phase of this case.

FACTUAL AND PROCEDURAL HISTORY

A. Background

Because of the limited purpose for which the Court has directed this briefing, Plaintiffs set out below only a brief summary of relevant facts. For a full factual recitation of the New York State Teacher Certification Examinations and a detailed description of the Core Battery and LAST, Plaintiffs respectfully refer this Court to their Proposed Findings of Fact and Judge Motley's Findings.⁷

The plaintiff class is comprised of African-American and Latino teachers who were prevented from, or delayed in, becoming full-time New York City public school teachers because of the requirement that they obtain a passing score on the Tests. Other than their failure to pass these examinations, Plaintiffs had met all other requirements to become certified as New

⁷ See Sohn Aff., Remand Exh. B (May 28, 2003 Plaintiffs' Proposed Findings of Fact and Conclusions of Law ("Plaintiffs' Proposed Findings")) at 12-13; Findings, ¶¶ 1-44.

York City public school teachers, including, but not limited to: (1) obtaining a bachelor's degree from a four-year college accredited by the New York State Education Department; (2) obtaining a masters degree from a similarly accredited institution; (3) earning a passing score on the State-mandated Assessment of Teaching Skills – Written (ATS-W), or an exemption from that requirement; (4) earning a passing score on the State-mandated content-specialty tests for the subjects that the individual teaches; and (5) completing a student-teaching period.⁸

SED and BOE used the Tests during different periods; the Core Battery from September 1984 through 1993, and the LAST from 1993 through the trial of this matter in 2002-03. At all times, the Tests were meant to function as certification exams that tested applicants' liberal arts and sciences knowledge. Teachers who failed to pass the Tests suffered significant consequences, including having their then-operative teaching credential revoked and being demoted to substitute teacher status with attendant decreases in salary of \$2,000 to \$20,000 per year. Despite this "demotion," in many cases, these teachers remained in the same classrooms teaching the same students, albeit at dramatically reduced compensation and benefit levels. In addition, after 1990, failure to pass these tests meant a teacher was ineligible for tenure and thus could not accrue seniority and advance up the salary scale. Despite the fact that both categories of teachers continued to teach students, there are significant differences between the employment rights and benefits of certified (those who passed the tests) and substitute (those who did not) teachers. Certified teachers have priority for placement in the City School District, greater system-wide seniority, and the right to secure another full-time position in the district if her

⁸ See Tr. 834:4-835:2; 884:17-887:8.

position is eliminated.⁹ The failure to properly validate the examinations and their misuse caused significant harm to numerous highly qualified, potential full-time public school teachers.

B. Standard Of Job Relatedness

On appeal, the Second Circuit established the framework within which the parties must work on remand. Specifically, the Second Circuit held that BOE can only prevail if it establishes the formal validity of the LAST under *Guardians*.¹⁰ The Second Circuit noted that although the terms “job relatedness,” “manifest relationship,” and “business necessity” have been used interchangeably by courts,

the basic rule has always been that “discriminatory tests are impermissible unless shown, by professionally acceptable methods, to be predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated.”¹¹

The Second Circuit distinguished *Watson v. Fort Worth Bank & Trust*¹² and the related case of *New York City Transit Authority v. Beazer*,¹³ relied on by Judge Motley, “as a subset of disparate impact cases in which the job relatedness of an employment practice is so patent that formal validation is unnecessary.”¹⁴ The Second Circuit concluded “[n]either of those outlying examples are relevant here.”¹⁵

⁹ See generally, Findings, ¶¶ 59-62.

¹⁰ *Gulino*, 460 F.3d at 385.

¹¹ *Id.* at 382-83 (citing *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 431 (1975)).

¹² 487 U.S. 977 (1988).

¹³ 440 U.S. 568 (1979).

¹⁴ *Gulino*, 460 F.3d at 386.

¹⁵ *Id.*

The Second Circuit also discussed the evidence necessary to establish whether *Guardians* has been satisfied, explaining that while expert testimony is helpful, “courts also need clearly established guideposts against which the reliability of the expert testimony can be evaluated.”¹⁶ For such guideposts, the Second Circuit held that “thirty-five years of using [the Equal Employment Opportunity Commission’s] Guidelines makes them the primary yardstick by which we measure defendants’ attempt to validate the LAST.”¹⁷

Thus, following the Second Circuit’s remand it is clear that BOE can only be successful if it establishes formal validity in accordance with *Guardians* and presents data and testimony that complies with professional standards at least as stringent as the Guidelines.

C. Pervasive Lack Of Documentation And Well-Established Record

In addition to the legal framework to be applied on remand, the Second Circuit defined the scope of evidence BOE may use to establish to validity. The Second Circuit affirmed Judge Motley’s conclusion that there is “insufficient documentation” to determine the “validity of the LAST.”¹⁸ The Second Circuit also concluded, however, that this “pervasive lack of documentation” was not necessarily fatal to formal validation. While “vague and unsubstantiated hearsay” is insufficient to establish job relatedness, “first-hand accounts of those involved in the test validation process” as well as experts “may be sufficient, in some

¹⁶ *Id.* at 383.

¹⁷ *Id.* at 384. In addition to the Equal Employment Opportunity Commission’s Guidelines (the “Guidelines”), Plaintiffs’ and Defendant’s experts agreed at trial that other professionally accepted standards relevant to test validation include the American Psychological Association Standards for Educational and Psychological Testing (the “Standards”) (see SED Exh. 122) and the National Research Council report “Testing Teacher Candidates” (see Pl. Exh. 253).

¹⁸ See *Gulino*, 460 F.3d at 387 (noting the “lamentable failure” of NES and SED to preserve documents “has considerably complicated an already difficult case”); Findings, ¶ 153.

circumstances” to establish formal validity.¹⁹ The Second Circuit did not opine that this case was such a “circumstance,” only that the possibility existed.

Thus, where, as here, a defendant is unable to meet the rigorous evidentiary burdens under the Guidelines through documentation,²⁰ a defendant may attempt to meet the same rigorous burdens through “first-hand” accounts of fact witnesses and experts. The Second Circuit was careful to explain, however, that this reprieve in no way lessened the evidentiary burdens of BOE nor “minimized” the “heavy burden that defendant bears in seeking to validate a test” without documentation.²¹

While the Second Circuit granted BOE this second chance, it did not send the parties back to square one. Instead, by affirming the “pervasive lack of documentation” in the record, the Court largely affirmed the factual findings of Judge Motley.²² The Second Circuit reversed as clearly erroneous only Judge Motley’s findings regarding the importance of the essay section of the LAST.²³ The remainder of Judge Motley’s findings are law of the case and controlling here.²⁴

¹⁹ *Gulino*, 460 F.3d at 388.

²⁰ See 29 C.F.R. 1607.15(A)(3) (“Where a total selection process has an adverse impact . . . the user should maintain and have available for each component of that process which has an adverse impact . . . [d]ocumentation evidence showing content validity . . . This evidence should be compiled in a reasonably complete and organized manner to permit direct evaluation of the validity of the selection procedure.”).

²¹ *Gulino*, 460 F.3d at 388.

²² *Id.* at 388-89 (upholding finding of a pervasive lack of documentation regarding the LAST).

²³ *Id.* at 386-87.

²⁴ See *U.S. v. Ben Zvi*, 242 F.3d 89, 95 (2d Cir. 2001) (holding mandate rule bars “relitigation of issues expressly or impliedly decided by the appellate court”); *County of Suffolk v. Stone & Webster Eng’g Corp.*, 106 F.3d 1112 (2d Cir. 1997) (“a decision made at a previous stage of litigation, which could have been challenged in the ensuing appeal but was not, becomes the law of the case; the parties are deemed to have waived the right to challenge that decision”).

ARGUMENT

I. JUDGE MOTLEY CORRECTLY CONCLUDED THAT THE LAST DOES NOT MEET THE REQUIREMENTS OF *GUARDIANS*

The field of test construction, “psychometrics,” is a scientific endeavor. As the APA Standards state, “[t]he process of validation involve[s] accumulating evidence to provide a sound scientific basis for the proposed score interpretations . . . [a]s in all scientific endeavors, the quality of the evidence is primary.”²⁵ Under the law of Title VII, face validity is not a substitute for legal proof of job relatedness.²⁶ Thus, to meet professional standards, BOE must present more than “vague and unsubstantiated hearsay,” rather, it must present evidence of a verifiable and repeatable process through which the assumptions and conclusions of the LAST can be tested by an independent fact-finder.

Moreover, it is irrelevant whether the LAST was promulgated pursuant to New York State’s police powers.²⁷ Although this argument has shifted over time and its place in the present context is not explained by BOE, the Second Circuit has conclusively established both the legal framework for determining the validity of the LAST and that BOE may be held liable under Title VII.²⁸ Regardless of how the LAST is characterized, job relatedness can only be

²⁵ SED Exh. 122 at 9-11; *see also id.* Standard 3.1 (“Tests and testing programs should be developed on a sound scientific basis.”); 29 C.F.R. § 1607.16(D) (content validity “[d]emonstrated by *data* showing that the content of a selection procedure is representative of important aspects of performance on the job”) (emphasis added).

²⁶ *NAACP v. Town of East Haven*, 70 F.3d 219, 223-24 (2d Cir. 1995) (finding error where court “uncritically accepted” defendant’s actions because it appeared logical that interest of public safety was served, noting that “[a]n assumption cannot pass for a factual finding”); *Berkman v. City of N.Y.*, 536 F. Supp. 177, 207 (E.D.N.Y. 1982) (holding that “while the concept of face validity may have some public relations role . . . , face validity is no substitute for validation based on concrete analysis either of the job’s requirements or the test’s results”), *aff’d*, 705 F.2d 584 (2d Cir. 1983).

²⁷ *See Sohn Aff.*, Remand Exh. L (November 9, 2009 Joint Letter to Honorable Judge Kimba Wood) at 7-8 (Defendant argued that “[t]he Court’s determinations as to the job relatedness of the LAST in the remand proceedings must . . . take into account the nature and purpose of the State’s police power licensing function as described by the Second Circuit”).

²⁸ *See Gulino*, 640 F.3d at 381 (“we affirm the district court’s decision that BOE is an ‘employer’ subject to liability under Title VII”).

established by valid, psychometric evidence. Because such evidence does not exist in the record, BOE cannot carry its burden under *Guardians* or Title VII.

A. **NES Did Not Perform A Suitable Job Analysis**

The Second Circuit has explained that a proper job analysis is one in which the employer breaks the job down into its component tasks, further breaks these tasks down into a set of component skills, and then determines the relative importance of these skills and the degree of proficiency required in each.²⁹ A job analysis cannot be established through “the general reputation of a test . . . its author . . . or casual reports of its validity . . . in lieu of evidence of validity.”³⁰

Here, as found by Judge Motley, the evidence established that no one conducted a proper analysis to identify the knowledge, skills, and abilities required to perform the job of New York public school teacher:

As shown below, neither NES nor SED properly identified the specific knowledge, skills and abilities that are critical to job performance and are likely to be present among the pool of people who meet the minimum qualifications of the job of New York public school teacher. Instead, NES presented test development committee members with test frameworks and a set of objectives that were never connected to the particular knowledge, skills, and abilities needed for successful performance of the job of New York public school teacher.³¹

According to NES, “the objectives and frameworks [of the LAST] were to be based on materials to be collected from the field, on teacher curricula from colleges and New York public schools,

²⁹ *Guardians* 633 F.2d at 242; *see also* Findings, ¶ 122 (noting the Guidelines require “focus on the work behavior(s) and the tasks associated with them. If the work behavior(s) are not observable, the job analysis should identify and analyze those aspects of the behavior(s) that can be observed and the observed work products. The work behavior(s) selected for measurement should be critical work behavior(s) and/or important work behavior(s) constituting most of the job”).

³⁰ 29 C.F.R. § 1607.9(A).

³¹ Findings, ¶ 121.

and from interviews to be conducted with Deans from selected colleges.”³² Although this two-part process was fundamental to identifying the KSAs that would be tested by the objectives, Judge Motley conclusively held “[t]here is no documentation describing any of the college materials collected or any information concerning the interviews.”³³ Before the Second Circuit, SED attempted to portray the collection and review of college material as organized and thorough, but no witness cited by the SED could recall what specific materials were collected or the substance of those materials.³⁴ No other witnesses could recall the particulars of the document collection or review process either.³⁵

The record similarly lacks any specific testimony regarding contemplated interviews with New York State public school teachers and educational experts. The Findings establish that the “only information concerning these interviews shows that a total of eight people were interviewed” and “[t]here is no information as to how the interviewees were selected, who they were, or what they said” nor whether the interview forms “still exist.”³⁶ Witnesses at trial had no

³² Findings, ¶ 125.

³³ *Id.*

³⁴ See Sohn Aff., Remand Exh. I (Second Circuit Brief for Defendant-Appellee The New York State Education Department (“SED Appellate Opposition Brief”)) at 16-17, citing Tr. 398:4-398:19 (witness recalled “collection of documents, materials and information” from “teacher education institutions,” but did not recall the “specifics” of the materials); 1125:18-1126:10 (witness testified that there were “offices filled” with materials collected from teacher colleges, but failed to identify any of this material or its specific source); 1126:16-1127:2 (testifying that “student curricular materials” were collected, but again failing to describe or identify materials with specificity); 1772:16-1772:23 (testifying that “course syllabi . . . college catalogues . . . [and] textbooks” collected, but failing to identify any materials individually).

³⁵ See, e.g., Tr. 1847:23-1848:4 (Stephen Weiss, an editorial director for NES, testified he did not know if there was “anyplace [one] could go to determine what material [was] collected.”); 396:20-397:2, 398:20-399:1, 400:21-401:5, 405:6-12 (Scott Elliot, the Vice President of Client Services for NES, testified he did not recall the process of gathering documentation for the LAST objectives, explaining “[i]t is mainly an issue of ten years lapse.” Dr. Elliot further testified he had no recollection who at NES collected New York college curricula, what specific reports were collected as part of the literature review, or the names of New York colleges that were contacted to collect materials and whose responsibility it was to contact those colleges.).

³⁶ Findings, ¶ 127 (citations omitted).

specific recollection of who was interviewed or how the interviewees contributed to the objectives.³⁷ Moreover, testimony cited by the SED before the Second Circuit in support of the interview process referred only in vague terms to the fact that interviews were conducted.³⁸

While NES represented that it had consulted with education experts, Judge Motley concluded that “no [expert] has been identified nor has any of their work been described other than in extremely general terms.”³⁹ Similarly, no witness was able to recall the identity of deans or faculty administrators interviewed or how they contributed to the LAST.⁴⁰ Thus, BOE can only rely on “casual reports” and “unsubstantiated hearsay” instead of verifiable evidence regarding who was interviewed, what they said, and how these interviews impacted test development, if at all. This does not meet professional standards.

³⁷ Tr. 416:3-6; 416:21-417:7 (Dr. Elliot testified he could not recall any interviewee by name, nor could he confirm that those interviewed represented a diverse population); Tr. 320:8-15 (Dr. Freeborne admitted that the eight interviews conducted were insufficient to account for geographically diverse points of view in New York); Tr. 799:4-7 (Ms. Hunsberger testified that she “certainly wouldn’t remember” who the interviewees were and instead referred the Court to NES); Tr. 1868:9-1869:6 (Mr. Weiss was similarly unable to recall what teachers were interviewed, how teachers were selected to be interviewed or what the results of the interviews were).

³⁸ SED Appellate Opposition Brief at 16-17, citing Tr. 398 (referring only to “discussions with educators”); 405 (“Q. And what occurred prior to – with respect to contacting teachers in the field, what was done prior to the drafting of the framework? A. I can’t recall the specifics, again, ten years later.”); 797-98 (testifying only that a teacher interview plan existed and information was supposed to be gathered); 810-11 (testifying that eight teacher interviews were conducted, but witness unaware of how any “fact-finder” could determine “what suggestions were made by these teachers”); 1125-26 (testimony refers only to collection of written materials in general); 1128 (testifying that a “number of folks were talked to,” but failing to identify any individual or what they said); 1777-80 (witness had no knowledge of who conducted interviews or who was interviewed, but testified about general types of information that were collected, however noting that he was unaware whether notes of the interviews still existed); 1849 (testifying that deans and faculty administrators were interviewed, but admitting he had no knowledge who was interviewed, who did the interviewing, or any specific recollection what was said in the interviews); 2834-36 (testifying that interviews did occur, but not offering any personal recollection of those interviewed or results from interviews).

³⁹ Findings, ¶ 133 (“Scott Elliot, a project director and later the Vice President of Client Services for NES during the relevant time periods, could recall neither the name of any content expert nor any records that would identify them or what they did. Tr. at 394-396. The most information that Stephen Weiss, an editorial director for NES, could remember were the names of two individuals who ‘commented on the content that we had to find for the various objectives.’ Tr. at 1585-26. Mr. Weiss testified that changes in the draft objectives were made by ‘national consultants,’ but he did not identify who they were or how anyone could determine what changes were actually made. Tr. at 1861-62.”).

⁴⁰ See, e.g., Tr. 1849:7-1849:10 (“Q. Is there a list of [deans and administrators] that were interviewed? A. I don’t know.”).

Finally, in addition to an inability to provide evidence of the process, Judge Motley found that “the retained documents indicate that both the frameworks and the objectives were drafted *before* the collection of any curriculum materials or the holding of any interview.”⁴¹ Thus, the information BOE alleged was collected to draft the objectives could not in fact have formed the basis of the KSAs which the LAST was meant to test.

On appeal, SED attempted to downplay this dramatic finding by claiming that formal validation had been established in other cases even when no job analysis was conducted before the test was given.⁴² To the contrary, the court in *AMAE* held that several job analysis studies *had* been conducted prior to the examination in question, which comported with the applicable professional standards.⁴³ Moreover, the *AMAE* court found that because the test at issue was an “off-the-shelf” test, it was unsurprising an employer would have to conduct validation studies after the test was created.⁴⁴ Here, BOE cannot present evidence that any professionally acceptable job analysis was conducted prior to drafting the framework and objectives.

It is axiomatic that a test cannot accurately predict whether an applicant will perform a job with minimal competence unless the KSAs required to perform that job with that competence have been previously identified. Before the Second Circuit, SED argued in a footnote that NES did not need to formulate KSAs for the LAST because “such KSAs were already reflected in the

⁴¹ Findings, ¶ 126 (emphasis added) (holding that a draft of the frameworks and objectives was provided to the SED on March 12, 1991, but that the “Plan for the Collection of College and University Curriculum Materials” was initially approved afterwards on March 18, 1991).

⁴² See SED Appellate Opposition Brief at 85, citing *AMAE v. California*, 937 F. Supp. 1397, 1419 (N.D. Cal. 1996), *aff’d*, 231 F.3d 572 (9th Cir. 2000).

⁴³ *AMAE*, 937 F. Supp. at 1419 (“Although Dr. Lundquist’s was the first *formal* job analysis, it is simply not true that no analysis of the relationship between the CBEST skills and the job of teaching was done before her study in 1995.”).

⁴⁴ *Id.*

liberal arts and sciences course work requirements.”⁴⁵ This argument, however, is circular and misguided. Because no suitable job analysis was completed, there is no evidence as to what KSAs a teacher must possess from a liberal arts education or elsewhere to be minimally competent. Here, both the documentation and testimony in the record are insufficient to identify “important elements of work behavior” “by professionally acceptable methods” and thus BOE cannot carry its burden.⁴⁶

B. NES Did Not Exercise Reasonable Competence In Test Construction

Beyond the fundamental building blocks of a valid and thorough job analysis, BOE also bears the burden of establishing that the processes followed by NES in constructing the LAST met professionally acceptable standards.⁴⁷ This BOE cannot do.

1. *The Composition And Work Of The Content Advisory Committee And The Bias Review Committee Were Insufficient*

A critical component to the construction of the LAST contemplated by NES was the work of the CAC and BRC. As SED argued before the Second Circuit, the purpose of these committees was to review the LAST and eliminate any bias.⁴⁸ Despite this critical function of test construction, the record lacks sufficient evidence to determine how members of these committees were selected or what their ultimate influence was on the LAST. These failings violate applicable professional standards.⁴⁹

⁴⁵ SED Appellate Opposition Brief at 80 n.21.

⁴⁶ See *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 431 (1975).

⁴⁷ *Guardians*, 630 F.2d at 94-95; *Gulino*, 460 F.3d at 384-85.

⁴⁸ SED Appellate Opposition Brief at 15-16.

⁴⁹ 29 C.F.R. § 1607.9(A) (“the general reputation of a test . . . its author . . . or casual reports of its validity [shall not] be accepted in lieu of evidence of validity.”); SED Ex. 122, APA Standard 3.5 (“When outside experts are used, their work should be documented . . . The qualifications, relevant experiences and demographic characteristics of expert judges should also be documented.”).

First, despite SED's argument in the Second Circuit that the CAC and BRC committee members were selected based on recommendations from deans of teacher colleges, superintendents, and principals,⁵⁰ Judge Motley found "that there would be no way now to review who was considered, accepted or rejected"⁵¹ for these committees, including "whether people were passed up who were more qualified."⁵² In addition, although the solicitations for nominees resulted in a "batch of documents that was 'three feet high,'" Judge Motley found "these documents were destroyed" in contravention of professional standards.⁵³

Second, Judge Motley concluded that while all BRC members were supposed to be "experienced in analyzing for test bias," this "requirement was subsequently dropped."⁵⁴ Instead, "the application/nomination form for committee members contains no reference to experience in analyzing for test-bias."⁵⁵ The only training SED claims that BRC members received was one presentation based on a brief outline of potential sources of bias.⁵⁶

In addition to the other deficiencies of the CAC and BRC, the only evidence of work conducted by these committees does not meet professionally accepted standards. The Findings establish that "[a]lthough notes were made and kept of discussions during committee meetings, only 'summary notes' on the 'consensus' of committee members were retained."⁵⁷ This type of

⁵⁰ SED Appellate Opposition Brief at 15.

⁵¹ Findings, ¶ 130.

⁵² Tr. 793:6-10.

⁵³ Findings, ¶ 130.

⁵⁴ Findings, ¶ 131.

⁵⁵ *Id.*

⁵⁶ SED Appellate Opposition Brief at 15.

⁵⁷ Findings, ¶ 131.

anecdotal evidence does not meet the requirements of Title VII validation.⁵⁸ The SED argued on appeal that the CAC and BRC conducted a review of the framework and objectives “lasting several days” and made revisions as necessary.⁵⁹ The testimony cited by SED, however, discusses only that revisions could have been made,⁶⁰ while no witnesses at trial were able to recall any specific revisions that were in fact made.⁶¹

Finally, it is not possible to discern the work of the CAC or BRC by the final product. As the Findings established, “[w]ith respect to the initial twenty-one test objectives and frameworks, only minor changes were made by the committees. The frameworks ultimately approved for use were substantially and substantively identical to the initial draft presented by NES in March, 1991.”⁶² Thus, not only was the work of the CAC and BRC not documented, but no witness at trial was able to offer any credible recollection of the work performed.

2. *The Work Of The Equity Advisory Board Was Insufficient*

NES represented that the EAB “generally reviews materials for inclusiveness and potential bias” and “has been proven invaluable at detecting potential problems for Blacks, Latinos, women, Native Americas, Asian-Americas, and individuals with handicapping conditions.”⁶³ On appeal, SED similarly characterized the EAB as “comprised of minority

⁵⁸ See *Williams v. Ford Motor Co.*, 187 F.3d 533, 540-41 (“under no circumstances will the general reputation of a test . . . its author . . . or casual reports of its validity be accepted in lieu of evidence of validity”); 29 C.F.R. § 1607.9(A).

⁵⁹ SED Appellate Opposition Brief at 17, 85.

⁶⁰ See, e.g., Tr. 283-84 (discussing committee selection process); 780-88 (witness testifies that LAST items were reviewed, but does not offer testimony about specific conclusions); 1129 (noting only that CAC had authority to change frameworks and objectives if necessary); 1132 (noting that CAC had authority to review “draft set of materials” and “revise it or delete it or add to it”); 2890-2903 (describing process of BRC in general terms).

⁶¹ See, e.g., Tr. 787:13-24; Findings, ¶ 132.

⁶² Findings, ¶ 132 (citations omitted).

⁶³ Findings, ¶ 134.

faculty members from colleges in the Amherst, Massachusetts area . . . which also reviewed the test at various stages of development.”⁶⁴ The testimony cited by the SED in its Second Circuit briefing, however, referred only to the conduct of the EAB in vague terms,⁶⁵ while the record at trial established that both the members and work of the EAB are a mystery.

The Findings establish that “no one from NES or SED who worked on the development of the LAST could identify a single person from the EAB with whom he or she met, or who did any work on the LAST” with the exception of Dr. Paula Nassif, who was able to “name only one current member of the EAB and no members from previous boards.”⁶⁶ Jeanne Clayton, who was identified as the employee most “likely” to have information on the EAB, “was unable to identify anyone on the EAB involved” with the LAST.⁶⁷

In addition, Judge Motley concluded that “[n]o one from NES or SED could articulate any specific action taken by the EAB beyond a vague suggestion to avoid portray [sic] Native Americas as victims in the test items.”⁶⁸ “No other witness could recall any other suggestion made by the EAB.”⁶⁹

⁶⁴ SED Appellate Opposition Brief at 15-16.

⁶⁵ Tr. 788-89 (witness testifies that EAB exists but is unaware whether any documents reflect work of EAB); 1204 (claiming EAB contained cultural bias experts, but failing to name any); 1780-81 (stated that EAB was comprised of experts, but failed to name any individual or their qualifications); 1786-93 (testifying that EAB exists, but unable to identify what bias training they received, who was on the EAB, or what comments were incorporated into the draft frameworks and objectives); 1850-51 (noting that draft objectives were reviewed at the March 12, 1991 meeting); 1860-61 (recalling only that EAB cautioned against portraying Native Americans as victims); 2683-85 (testifying that EAB reviewed LAST objectives, but admitting that she is unable to recall any recommendations on specific objectives).

⁶⁶ Findings, ¶ 135.

⁶⁷ *Id.*

⁶⁸ Findings, ¶ 136.

⁶⁹ *Id.*

The failure to document the identity or expertise of the members of the EAB violates professional standards.⁷⁰ In addition, the inability of anyone to identify the work of the EAB on the LAST undermines any suggestion of validity.⁷¹

3. *Item Drafting For The LAST Did Not Meet Professionally Acceptable Standards*

After the objectives and frameworks had been drafted and reviewed, individual test questions, or “items,” were drafted. On appeal, SED attempted to create the illusion that specific testimony regarding the iterative steps of item development existed, however, a closer look at the cited testimony reveals that no witness could offer specific recollections of when or how items were written or changed.⁷² This lack of any detailed evidence fails to meet professionally acceptable standards for test item development.⁷³

Judge Motley found that no “witness could identify any person involved in drafting the items” and “that there is no way of determining who wrote items and what they were assigned to write.”⁷⁴ Furthermore, neither BOE nor SED at trial was able to offer any “documentation [of]

⁷⁰ See 29 C.F.R. § 1607.9(A); SED Exh. 122, Standard 3.5.

⁷¹ *Id.*

⁷² See, e.g., SED Appellate Opposition Brief at 19, citing Tr. 1816-21 (witness testified that item try-outs conducted, but cannot recall comments from participants, was reviewed by content experts, but cannot identify who the experts were or what they did, that items underwent various revisions, but cannot identify why an item was revised, and the EAB reviewed draft items, but is unable to identify any actions taken by EAB); 2685-87 (noting that EAB, CAC, and BRC reviewed test items, but witness is unable to identify any work done on the items by these three committees).

⁷³ See SED Exh. 122, APA Standard 3.7 (requiring full documentation of the “procedures used to develop, review and try out items, and to select items from the item pool”); Standard 3.8 (item try-outs “should be documented” and the “samples should be as representative as possible of the populations for which the test is intended”); 29 C.F.R. § 1607.9(A).

⁷⁴ Findings, ¶ 139.

the background, expertise or experience of any item writer.”⁷⁵ The inability to offer any information on the background of experts used to draft the test items violates the Guidelines.

Perhaps more significantly, the Findings establish that there is “no way of knowing what source material was used to create any test item,” nor is there any way “of determining what changes were made to any of the draft items.”⁷⁶ Although NES represented to SED that it would use a form for each test item that would record the “author, date, objective, topic, item number and source, and included a large space for the drafting of each item,” no witness at trial recalled whether such a form was actually used.⁷⁷ Finally, although “item tryouts” were allegedly conducted and changes made based on these tryouts, “there is no documentation of this activity and no way of determining what was done in this regard.”⁷⁸

The utter lack of evidence regarding the specific steps taken to develop and try-out test items violates both APA Standard 3.7 and 3.8, as well as the Guidelines.

4. Pilot Testing Of The LAST Did Not Meet Professionally Acceptable Standards

After test items were developed, pilot tests were conducted at twenty-two teacher colleges in New York. These pilot tests were not full-length LAST examinations, but abridged versions containing only thirty-six questions instead of the eighty questions (sixty-four scored) in a full-length examination. These pilot administrations did not meet professionally acceptable standards of test construction because the population used for these pilot tests was not

⁷⁵ *Id.*

⁷⁶ Findings, ¶ 140; *see also* Tr. 1870:18-1871:2 (“Is there anything I could do now to determine what changes were made [to the draft items]? A. No.”).

⁷⁷ Findings, ¶ 140.

⁷⁸ *Id.*

representative of the actual test population and the shortened-sample tests were not validated against the full-length exams.⁷⁹

Individuals used to test the exams were current college students preparing to be teachers and did not include any in-service teachers. In addition, the vast majority of those who took the pilot tests had no teaching experience whatsoever.⁸⁰ According to NES, students were used for the pilot test because they could be compelled to take the test without being paid and because most of the people who would be taking the test would be in college.⁸¹ NES did nothing, however, to determine whether there was a statistically significant difference in pass rates between college students and working teachers.⁸² In addition, NES did not describe the criteria used for determining which items were psychometrically sound and which were flawed, thus providing no basis for which questions should be deleted or retained.⁸³

⁷⁹ See SED Exh. 122, APA Standard 3.8 (“When item tryouts or field tests are conducted, the procedures used to select the samples(s) of test takers for items tryouts and the resulting characteristics of the sample(s) should be documented. When appropriate, the sample(s) should be as representative as possible of the population(s) for which the test is intended.”); *id.*, Standard 3.16 (“If a short form of a test is prepared by reducing the number of items or organizing portions of a test into separate form, empirical data or a theoretical rationale should be provided to estimate the reliability of each short form and its correction with the standard form.”).

⁸⁰ Tr. 828:3-829:15; Pl. Exh. 153.

⁸¹ Tr. 1166:21-1167:12.

⁸² Tr. 1205:5-9.

⁸³ See Pl. Exh. 242 (Rebuttal Expert Report of Dr. Frank Landy) at 25.

Regarding the short-form nature of the pilot tests, Plaintiffs' expert, Dr. Frank Landy,⁸⁴ found several elements of the short-form tests did not meet professionally acceptable standards. First, NES did nothing to connect performance on the short-form test to predicted performance on a full-length examination.⁸⁵ Second, because ten different test forms were used in the pilot, only small statistical results were able to be gathered for any one question and NES did not establish these statistics were representative of all test-takers.⁸⁶ This concern is augmented by the fact that the items on the ten test forms did not perform in the same manner. Thus, as Dr. Landy concluded, "NES could not reliably predict the performance of the operational tests."⁸⁷

5. Defendant Violated Professional Standards By Failing To Conduct Continuing Procedures To Eliminate Bias

Before the Second Circuit, SED argued that "statistical analyses about candidates' performance" are conducted by NES after each "test administration" to determine "the performance of different racial and ethnic groups."⁸⁸ What SED did not mention, and Judge Motley concluded, was that "[t]he disparity in passing rates on the LAST has increased over

⁸⁴ Dr. Frank J. Landy obtained a Masters degree and Ph.D. in industrial organizational psychology and psychometrics from Bowling Green State University in 1966 and 1969 respectively. Dr. Landy was a Professor of Psychology at Penn State University from 1969 until 1994 and held the position of Professor Emeritus from 1994 through the date of trial. Dr. Landy was also a visiting Professor of Psychology at Stanford University and a visiting Researcher in the Department of Psychology at the University of California, Berkeley. In addition to his academic appointments, Dr. Landy was the President of Landy, Jacobs and Associates, Inc. from 1984 through 1998 and then the CEO of SHL Landy Jacobs Litigation Support Group from 1998 through the date of trial. At these companies Dr. Landy was involved in the development of hundreds of employment-related tests as well as counseling numerous employers on selection and employment decisions. Dr. Landy had presented hundreds of lectures in the field of testing and written numerous articles and books. Dr. Landy had also testified and been qualified by Federal Courts as an expert witness. (*See generally*, Tr. 1262:9-1267:22; 1299:18-24; Pl. Exh. 241, Vita of Dr. Landy.) Dr. Landy passed away in January 2010.

⁸⁵ Pl. Exh. 242 at 26-27; SED Exh. 1 (Expert Report of Dr. Frank Landy) at 57.

⁸⁶ *Id.*

⁸⁷ *Id.*

⁸⁸ SED Appellate Opposition Brief at 25.

time,” yet “no differential item analysis was undertaken”⁸⁹ to assess the cause of the disparity in passing rates. Although NES reviewed “statistical analyses about candidates’ performance” after each test administration,⁹⁰ Ms. Clayton testified that she was not aware of any item that had been discarded after a review of test-takers’ performance.⁹¹ Judge Motley found this total failure to recognize and account for increasingly disparate results among minority applicants “violates APA Standard 7.3.”⁹²

This failure of NES was compounded by its failure to perform other processes to monitor the LAST for bias. For example, although NES was to convene an annual meeting of the Minority Review Committee, Judge Motley found “[t]here is no evidence [this] occurred.”⁹³ At the BRC meeting in 1993, the Committee reviewed statistics on the first administration of the test. Those results showed that of the 3,110 people who took the test, only 49 were African-American and only 85 were Latino.⁹⁴ After that meeting, the BRC recommended that there be annual meetings to review the cut scores.⁹⁵ SED rejected this recommendation.⁹⁶ There was a

⁸⁹ Findings, ¶ 142.

⁹⁰ SED Appellate Opposition Brief at 25.

⁹¹ Pl. Exh. 289 at 75-76, 83-84.

⁹² Findings ¶ 142; see SED Exh. 122, APA Standard 3.7 (“When credible research reports that differential item functioning exists across . . . racial/ethnic, cultural . . . groups in the population of test takers . . . test developers should conduct appropriate studies when feasible. Such research should seek to detect and eliminate aspects of test design, content, and format that might bias test scores for particular groups.”); *Fickling v. New York State Dep’t of Civil Serv.*, 909 F. Supp. 185, 191 (S.D.N.Y. 1995) (finding significant defendants’ failure “to conduct[] a predictive validity study during the 15 years that these examinations were used, despite their knowledge for over two years of the examinations’ disparate impact”).

⁹³ Findings, ¶ 142.

⁹⁴ Pl. Exh. 160.

⁹⁵ Pl. Exh. 173.

⁹⁶ Tr. 714:2-715:13.

second review in 1995.⁹⁷ At the conclusion of that meeting the BRC again recommended that NES consider an “annual cutscore review . . . [and] a review of data analysis that provides recommendations for improving minority performance”⁹⁸ There is no evidence of any further meetings of the BRC itself in which minority performance was reviewed for potential item bias or disparate impact. Since the initiation of the LAST, test statistics have been reviewed by the BRC on only two occasions – in 1993 and 1995.⁹⁹

6. The Defendant Presented No Evidence Of Purported Changes To The LAST Based On New Learning Standards In The Late 1990s

On appeal, SED claimed that from 1996 through 1998 the LAST was re-validated pursuant to new learning standards promulgated by the State of New York.¹⁰⁰ Judge Motley, however, conclusively found no “re-validation” was established. The Findings held that “there is no way of determining what was done [from 1996 through 1998], whether the items were changed or how any new items reflected the new learning standards.”¹⁰¹ Thus, even this supposed re-validation offers no support to BOE’s position.

7. Defendant’s Past Justification For The Poor Test Construction Of The LAST – That Students Of Certified Teachers Perform Better In The Classroom – Has No Basis In The Record

Lacking a psychometric basis to argue the LAST is valid, BOE and SED in prior proceedings attempted to imply that the LAST remains valid because teachers who pass the

⁹⁷ Pl. Exh. 289 at 91, 176.

⁹⁸ Pl. Exh. 174.

⁹⁹ Tr. 708:14-709:13; 2668:3-13; 2672:6-12; Pl. Exh. 289 at 74, 153.

¹⁰⁰ SED Appellate Opposition Brief at 28-29.

¹⁰¹ Findings, ¶ 141.

LAST are better teachers and perform better than those who did not pass.¹⁰² This argument finds no support in the record. No witness testified that any such study was undertaken and the two outside authorities relied on by BOE at trial said no such thing.¹⁰³

The Findings established that in the ten years SED administered the LAST prior to trial of this case in 2002, it never undertook a single “study correlating teachers’ performance on the LAST with the performance of students in the classroom.”¹⁰⁴ Jeanine Grinage, the acting Deputy Commissioner of Higher Education between February 1996 and May 1998, thought such a study “would be a ‘good idea’” and Commissioner of Education Richard Mills asked for such a study, but one has never been conducted.¹⁰⁵ Numerous additional witnesses testified at trial that no study exists that establishes students whose teachers pass the LAST perform better than students whose teachers did not pass the LAST.¹⁰⁶ Any suggestion to the contrary is simply not accurate.

C. The Content Of The LAST Is Not Related To The Content Of The Job

As this Court has stated, “[t]he crux of Title VII’s requirements is that ‘the content of the test must be related to the content of the job.’”¹⁰⁷ In order for the content of a test to be considered sufficiently related to the content of the job, the employer must demonstrate that the skills identified in the job analysis are “essential to effective job performance,” and that the test

¹⁰² See SED Appellate Opposition Brief at 36-37.

¹⁰³ See generally Sohn Aff., Remand Exh. E (June 11, 2003 Plaintiffs’ Reply Proposed Findings of Fact and Conclusions of Law) at 13-16.

¹⁰⁴ Findings, ¶ 151.

¹⁰⁵ *Id.*

¹⁰⁶ See Sohn Aff., Remand Exh. K (Second Circuit Reply Brief of Plaintiffs-Appellants-Cross-Appellees) at 6-8.

¹⁰⁷ *Cuesta v. NY State Office of Court Admin.*, 657 F. Supp. 1084, 1098 (S.D.N.Y. 1987) (quoting *Guardians*, 630 F.2d at 95).

directly measures an applicant's possession of those skills.¹⁰⁸ Here, since neither BOE, SED nor NES conducted a thorough and meaningful job analysis, or even a minimally sufficient job analysis, it is impossible to determine whether the content of the LAST is sufficiently related to the content of the job so as to justify its use.¹⁰⁹

In a proper test-construction process, a test developer would create a list of the important KSAs required for the job and then translate those KSAs into test questions.¹¹⁰ Here, because NES, SED, and BOE never created a list of important KSAs, the development of test questions was necessarily not related to any meaningful analysis of the job of a public school teacher.¹¹¹ Accordingly, BOE cannot meet its burden of proving that the content of the LAST is related to the content of the job.

Similarly, the LAST has no identifiable domain, which means that there is no identifiable universe of subject matter, the knowledge of which the test is designed to measure. An identifiable test domain is important because it specifies the body of knowledge from which the test is drawn, and because it informs applicants of the areas they need to study and where they can obtain the information that will be tested.¹¹² The LAST has no such domain. In fact, at trial, Dr. Mehrens testified that he did not know what areas applicants should study and suggested that

¹⁰⁸ See *Dothard v. Rawlinson*, 433 U.S. 321, 331 (1977); *Lanning v. SEPTA*, 181 F.3d 478, 486 (3d Cir. 1999).

¹⁰⁹ See *Vulcan Soc. of New York City Fire Dep't v. Civil Serv. Comm'n*, 490 F.2d 387, 395 (2d Cir. 1973) ("An examination has content validity if the content of the examination matches the content of the job. For a test to be content valid, the aptitudes and skills required for successful examination performance must be those aptitudes and skills required for successful job performance."); see generally, *Guardians*, 630 F.2d at 94-95; *Gulino*, 460 F.3d at 384-385.

¹¹⁰ See *Cuesta*, 657 F. Supp. at 1098 ("the professional translation of job analysis results into test items provides great assurance that the abilities tested for are actually required for the job.").

¹¹¹ See Findings, ¶ 121 ("neither NES nor SED properly identified the specific knowledge, skills and abilities that are critical to job performance").

¹¹² Tr. 1357:21-1358:14.

in order to determine what to study, an applicant should talk to experts in all the fields being tested. He did not dispute that the unlimited domain of the LAST could include every artistic work ever created and every book ever written.¹¹³

BOE and SED have previously contended that a “broad domain” of “liberal arts and sciences” is sufficiently defined and identifiable, providing the example of the “thousands” of colleges that offer degrees in liberal arts and sciences.¹¹⁴ To argue that every liberal arts and sciences college teaches the same curriculum, or every student who graduates has the same core knowledge, however, ignores the reality—and purpose—of such an education. SED also previously argued that Dr. Landy agreed the LAST contains only content in the statewide core curriculum and agreed certain skills are fundamental and require no analysis of domain.¹¹⁵ First, it is simply wrong that Dr. Landy agreed that the LAST contains only content in the statewide core curriculum.¹¹⁶ Second, while Dr. Landy did agree “basic mathematics,” “basic minimum reading skills,” and “basic minimum writing skills” should be something every teacher knows, the LAST tests areas far beyond these “basics” and the failure to define a domain beyond these basics violates professional standards.¹¹⁷ BOE’s own witnesses at trial disagreed as to the range

¹¹³ Tr. 3225:15-3228:5.

¹¹⁴ See, e.g., Sohn Aff., Remand Exh. F (June 11, 2003 SED Reply to Plaintiffs’ Proposed Findings of Fact) at 11, n. 17.

¹¹⁵ SED Appellate Opposition Brief at 37-38.

¹¹⁶ Tr. 1443:11-14 (“Q. Doctor, you do not disagree with the proposition that there is nothing on the LAST that is not contained in the New York State curricula? A. I don’t know. I haven’t seen the New York State curricula.”).

¹¹⁷ See also Tr. 1442:23-1443:10 (Q. Do you dispute that a public school teacher should possess some minimum levels of knowledge of literature and fine arts, mathematics, science and social studies? A. Those are unbounded domains with the exception of mathematics. I don’t disagree that there should be some minimal level and that we can all agree on what basic mathematics are. But with respect to those other three elements, their data analysis showed very clearly that they could not develop any measures of those things, that people could not respond to them as if they were unique things. And that conceptually we cannot even define what they are beyond the fact that there may be a course title in one university but not another to cover something like one of those.”).

of subjects included within the category of “liberal arts and sciences.”¹¹⁸ Without a defined domain for the content of the test, BOE cannot show that the content of the test is related to the content of the job.

D. The Content Of The LAST Is Not Representative Of The Content Of The Job

The fourth prong of the *Guardians* standard, regarding the representativeness of the test, has two components.¹¹⁹ First, the content of the test must be representative of the content of the job. Second, the procedure, or methodology, of the test “must be similar to the procedures required by the job itself.”¹²⁰ Although NES contemplated a job-analysis survey to solicit opinions regarding the representativeness of the objectives, the multiple failures of this survey render it impossible for BOE or SED to demonstrate that the LAST is representative of any job.

According to its Job Analysis Survey Plan, NES was supposed to select a random sample of up to 800 teachers and college faculty to complete a questionnaire rating the importance of twenty-one test objectives.¹²¹ There is no documentation as to how that sample was prepared and witnesses most likely to have knowledge could not explain how the sample was selected.¹²² In the questionnaire, participants were asked: “How important is the knowledge or skill described by this objective for performing the job of an educator in New York State public schools? 1) no importance, 2) little importance, 3) moderate importance, 4) great importance, 5)

¹¹⁸ See, e.g., Tr. 2982:18-2983:17; 3027:2-9; 3075:15-20; 3439:13-3441:8; 3442:11-3443:8; 3400:20-3401:7.

¹¹⁹ *Guardians*, 630 F.2d at 94-95.

¹²⁰ *U.S. v. Vulcan Soc., Inc.*, 637 F. Supp. 2d 77, 118 (E.D.N.Y. 2009).

¹²¹ Pl. Exh. 130.

¹²² See Pl. Exh. 289 at 129-30 (Deposition of Jeanne Clayton, Senior Area Director for NES); Tr. 442:5-23 (Testimony of Dr. Elliot).

very great importance.”¹²³ No definition of the term “importance” was given to survey participants and even Dr. Elliot acknowledged that the participants could have different contexts for judging importance.¹²⁴ Similarly, the Job Analysis Survey gave no definition for harm or otherwise indicated that the purpose of the test was to prevent harm.¹²⁵ In fact, there is no evidence that NES, SED, or BOE ever discussed the concept of harm, specified the harms to be protected against, or designed the LAST to ensure that harm was prevented.¹²⁶

Participants in the survey were also not informed by the questionnaire that their rating of the objectives determined not only the content of the test items, but also the relative weights to be given the test items.¹²⁷ There is no way to prove the representativeness of the LAST because there was never an empirical study performed supporting the weights accorded the various subjects on the tests. For example, Ms. Hunsberger testified that she was not aware of any empirical data supporting the decision to make the essay portion of the test count as twenty percent of the total score, nor could she explain the empirical basis for the weighting of items on the test other than that the number of questions was to be based on the number of objectives.¹²⁸ SED’s own expert testified that he would have weighted the LAST objectives differently if he had his “druthers.”¹²⁹ As a result, it cannot be demonstrated that the content of the LAST is representative of the content of the job.

¹²³ Pl. Exh. 132.

¹²⁴ Tr. 444:17-445:1.

¹²⁵ Tr. 3149:4-3150:3 (Testimony of Dr. Mehrens); Findings, ¶ 148.

¹²⁶ Tr. 1360:12-1361:1 (Testimony of Dr. Landy).

¹²⁷ Tr. 820:11-17 (Testimony of Edith Hunsberger, Supervisor of Education Programs).

¹²⁸ Tr. 817:8-819:13.

¹²⁹ Tr. 3158:116-17.

E. The Scoring System Of The LAST Does Not Distinguish Between Those Who Will And Will Not Cause Harm In The Classroom

As the *Guardians* Court stated, “[n]o matter how valid the exam, it is the cutoff score that ultimately determines whether a person passes or fails. A cutoff score unrelated to job performance may well lead to the rejection of applicants who were fully capable of performing the job. When a cutoff score unrelated to job performance produces disparate racial results, Title VII is violated.”¹³⁰

According to Plaintiffs’ and SED’s experts, the purpose of a “cut score” is to distinguish between applicants who are minimally competent and those who are incompetent.¹³¹ In order to determine a cut score, “there should generally be some independent basis for choosing the cutoff.”¹³² For example, an employer “might establish a valid cutoff score by using a professional estimate of the requisite ability levels, or, at the very least, by analyzing the test results to locate a logical ‘break-point’ in the distribution of scores.”¹³³ Here, as in *Guardians*, BOE and SED did neither.

To establish the LAST cut score, the CAC and BRC used the Modified Angoff Method.¹³⁴ This consisted of asking groups from the committees to review 80 questions from the first LAST and then answer the following question:

Imagine a hypothetical group of individuals who have the minimally acceptable level of knowledge and skills required to perform the job of an educator receiving a provisional certificate in

¹³⁰ *Guardians*, 630 F.2d at 105.

¹³¹ Findings, ¶ 152.

¹³² *Guardians*, 602 F.2d at 105.

¹³³ *Id.*

¹³⁴ Findings, ¶ 149.

New York State. What percent of this group would answer the item correctly?

No one from SED, NES, or the committees performed an Angoff analysis of the other 270 items in the 350 item bank.¹³⁵ Using this method, which, according to Dr. Landy's expert opinion is the least desirable method for setting a cut-score,¹³⁶ the CAC selected a cut score of 48 and the BRC selected a cut score of 38.¹³⁷ The State Commissioner approved a passing score of 38.¹³⁸ The passing score was later raised to 44, then lowered to 43, even though the passing score committee was provided statistics that showed a cutoff score of 44 would mean only 29% of African-American test-takers and only 26% of Latino test-takers would pass, as opposed to 78% of White test-takers.¹³⁹

Neither the acting Deputy Commissioner of Higher Education who recommended raising the cut score nor the Commissioner who followed the recommendation had, at the time of the recommendation, ever seen a complete copy of the LAST or knew of any study correlating teachers' performance on the LAST with the performance of students in the classroom.¹⁴⁰ In other words, the LAST cut score was raised without any justification. According to Ms. Torres, who sat on one of the committees that recommended raising the cut score on the LAST to 44, the purpose of the LAST cut score is not to distinguish between minimally competent and incompetent teachers.¹⁴¹ As stated above, however, this is precisely what the purpose of a cutoff

¹³⁵ *Id.*

¹³⁶ Tr. 1338:23-1339:19.

¹³⁷ Findings, ¶ 149.

¹³⁸ *Id.*

¹³⁹ Tr. 596:1-6; Pl. Exhs. 212, 214, 215, 217, 218.

¹⁴⁰ Findings, ¶ 151; Tr. 3339:19-3340:4.

¹⁴¹ Findings, ¶ 152.

score on a certification test should be. Accordingly, the scoring system of the LAST does not usefully select from among the applicants those who can better perform the job of public school teacher as *Guardians* requires.

In addition to the flawed methodology in choosing a cut score, there were also serious flaws in the scoring of the essay portion of the LAST. Because the scoring of an essay necessarily entails a subjective evaluation, test developers must ensure and document the consistency of evaluations made by different scorers.¹⁴² The Supreme Court has held that when there is no way to know precisely what criteria scorers considered, whether each scorer considered the same criteria, and whether the scorers applied a focused and stable body of criteria, there is no way to determine whether the criteria considered was sufficiently job-related.¹⁴³ Here, although NES made representations that it would develop a “statistical assessment of the reliability of the scoring process and of the consistency of the scores of the independent scorers,”¹⁴⁴ there is no evidence this was done. Without this evidence, BOE cannot meet its burden to demonstrate validity of the essay portion of the LAST.

For all these reasons, the development of the LAST failed to meet the minimum standards embodied by the Guidelines and *Guardians* and thus validity cannot be established. Accordingly, BOE should be found liable under Title VII.

¹⁴² See 29 C.F.R. § 1607.149(C)(5) (“appropriate statistical estimates should be made of the reliability of the selection procedure”).

¹⁴³ *Albemarle Paper Co.*, 422 U.S. at 433.

¹⁴⁴ Pl. Exh. 95.

II. THE LAST AND CORE BATTERY WERE MISUSED TO MAKE EMPLOYMENT DECISIONS REGARDING IN-SERVICE TEACHERS

A. The LAST

Regardless of whether the LAST was properly developed for teacher candidates about to enter the classroom—which it was not—that was its intended purpose.¹⁴⁵ Consequently, BOE misused the LAST when it administered the exam to in-service teachers and then used the results to make employment decisions about those teachers, including decisions about salaries, pensions, and seniority.¹⁴⁶ Controlling authorities unequivocally provide that the misuse of properly designed tests renders those tests invalid.¹⁴⁷ For example, APA Standard 6.3 provides that “[w]hen a test is to be used for a purpose for which it has not been previously validated or for which there is no supported claim of validity, the user is responsible for providing evidence of validity.”¹⁴⁸ Similarly, APA Standard 11.5 prohibits the use of a test for purposes not intended by the test developer.¹⁴⁹

Courts similarly prohibit test users such as BOE from using a test for a purpose other than that for which it was intended.¹⁵⁰ For example, in *Walston v. School Board* the Fourth Circuit enjoined the use of Core Battery scores for making employment decisions unless the test was

¹⁴⁵ Pl. Exh. 47; Pl. Exh. 287 at 143-49; Tr. 174:15-24.

¹⁴⁶ Plaintiffs’ Proposed Findings, ¶¶ 35-37.

¹⁴⁷ Pl. Exh. 241, at 36-37; SED Exh. 122 (1999 APA Standards) §§ 6.3, 11.3, 11.5.

¹⁴⁸ SED Exh. 122 § 6.3; Pl. Exh. 241 at 36.

¹⁴⁹ SED Exh. 122 § 11.5 (“[t]est users should be alert to the potential misinterpretation of test scores and to possible unintended consequences of test use; users should take steps to minimize or avoid foreseeable misinterpretations and unintended negative consequences”).

¹⁵⁰ See *Walston v. County School Bd. of Nansemond County, Virginia*, 492 F.2d 919, 927 (4th Cir. 1974) (holding that local school district misused the Core Battery when it based employment decisions on it).

specifically “validated for use for that purpose.”¹⁵¹ The entity using a test is responsible for proving the validity of the particular use.¹⁵²

Here, the LAST test developer, NES, consistently warned all local school districts in New York, in an annual “[i]nterpretive caution,” that “[i]ndividual examinee results . . . are for the purpose of New York State teacher certification only . . . [and] are NOT intended to be used for employment decisions, . . . or for any other purpose.”¹⁵³ The president of NES testified that in conversations with SED officials, he made it clear that the “LAST test is . . . not designed for the purpose” of cutting salaries or any other employment related purpose.¹⁵⁴ In 1997, because she was concerned that LAST scores were being misused, SED’s deputy commissioner of higher education circulated a memo to local school districts warning that the LAST was “not validated . . . for . . . purposes such as . . . employment screening.”¹⁵⁵

Despite these warnings, however, results on the LAST consistently were used to make decisions regarding the salaries, seniority, pensions, and other benefits of New York City teachers.¹⁵⁶ Failure to pass the LAST and demotion to substitute teacher status could mean a salary decrease of up to \$20,000 per year, loss of priority for placement in the City School

¹⁵¹ See *Walston v. School Bd.*, 566 F.2d 1201, 1204 (4th Cir. 1977); see also *York v. Alabama State Bd. of Ed.*, 581 F. Supp. 779 (M.D. Ala. 1983) (granting preliminary injunction prohibiting defendants from misusing the Core Battery).

¹⁵² See SED Exh. 221, 1999 APA § 6.3; see also *EEOC v. Atlas Paper Box Co.*, 868 F.2d 1487, 1490-91 (6th Cir. 1989); *Bernard v. Gulf Oil*, 841 F.2d 547, 567 (5th Cir. 1988).

¹⁵³ Pl. Exh. 47.

¹⁵⁴ Pl. Exh. 287 at 143-49.

¹⁵⁵ Pl. Exh. 46; Tr. 585:5-586:23.

¹⁵⁶ Plaintiffs’ Proposed Findings, ¶¶ 35-37.

District, loss of system-wide seniority, and other benefits enjoyed by certified teachers.¹⁵⁷ The consequences to in-service teachers of failing the Tests was severe.

Plaintiffs' expert, Dr. Landy, found that both the LAST and Core Battery (discussed below) were misused when BOE based employment decisions for in-service teachers on scores from these tests.¹⁵⁸ SED's expert did not disagree with Dr. Landy concerning misuse and did not offer *any* opinion on the issue. Moreover, BOE did not offer any expert testimony on this topic. Accordingly, BOE violated Title VII by using the LAST in a manner for which it was not intended or validated.

B. The Core Battery

1. *The Second Circuit Did Not Consider Whether The Core Battery Was Misused And Therefore The Issue Remains Open On Remand*

Although appealed by Plaintiffs, the Second Circuit did not address whether the Core Battery was misused and thus the issue may be decided by this Court on remand.¹⁵⁹ The mandate rule, as articulated by the Second Circuit, directs that "if an [appealed] issue was not part of the appellate decision, a trial court may consider the matter."¹⁶⁰ Furthermore, in considering whether an issue remains open on remand, the "court should look to both the specific dictates of the remand order as well as the broader spirit of the mandate."¹⁶¹ Here, the Second Circuit specifically directed that this Court may consider the issue of misuse on remand, but did

¹⁵⁷ See generally, Findings, ¶¶ 59-62.

¹⁵⁸ See Plaintiffs' Proposed Findings, ¶¶ 220-226; Pl. Exh. 241 at 36-37, 60; Tr. 1418: 17-20; 1420: 6-10; 1508:8-15.

¹⁵⁹ See Sohn Aff., Remand Exh. H (Second Circuit Brief of Plaintiffs-Appellants-Cross-Appellees) at 57-62 (arguing both Core Battery and LAST were misused when used to make employment decisions regarding in-service teachers).

¹⁶⁰ U.S. v. Minicone, 994 F.2d 86, 89 (2d Cir. 1993); see also Burrell v. U.S., 467 F.3d 160, 165 (2d Cir. 2006) (quoting *Minicone*).

¹⁶¹ Meacham v. Knolls Atomic Power Lab., No. 09-2037-CV, 2009 WL 4912205, at *1 (2d Cir. Dec. 21, 2009) (quoting U.S. v. Ben Zvi, 242 F.3d 89 (2d Cir. 2001)).

not address the Core Battery.¹⁶² In so doing, the mandate rule and the spirit of the remand order reasonably permit, if not compel, the Court to consider misuse of both the LAST and the Core Battery.¹⁶³

2. *The Core Battery Was Misused*

As with the LAST, the evidence at trial made plain that the Core Battery was intended to be used to make decisions regarding initial certification (*i.e.*, whether an individual should be allowed into the classroom) and not to make decisions regarding salaries, pensions, or seniority.¹⁶⁴ Educational Testing Service (“ETS”), the Core Battery’s developer, repeatedly expressed concern that the Core Battery not be misused for “career ladder” employment decisions in publications to its clients, including the State of New York.¹⁶⁵ The president of ETS also specifically warned about misuse of the Core Battery in making employment decisions.¹⁶⁶ The ETS psychometrician responsible for developing the Core Battery testified at trial that ETS did not allow the exam to be used by employers to reduce salaries, pension rights, and seniority.¹⁶⁷

Nevertheless, BOE in fact used the Core Battery for just such purposes.¹⁶⁸ Both SED and BOE officials understood this prohibition and acknowledged that the test was being misused when its results were the sole basis for reducing the salaries, pensions, and seniority rights of

¹⁶² *Gulino*, 460 F.3d at 370, n. 9.

¹⁶³ See *Minicone*, 994 F.2d at 89; *Burrell*, 467 F.3d at 165; *Meacham*, 2009 WL 4912205, at *1.

¹⁶⁴ Pl. Exhs. 78; 79; 80.

¹⁶⁵ Pl. Exhs. 78; 79, at 15-24; 80.

¹⁶⁶ Pl. Exh. 254.

¹⁶⁷ Tr. 1967.

¹⁶⁸ Plaintiffs’ Proposed Findings, ¶¶ 35-37.

class members.¹⁶⁹ In fact, one BOE official testified that they knew “they were pushing the envelope” on the use of the Core Battery to make employment decisions.¹⁷⁰ Plaintiffs’ expert further established that such tests were not validated for making employment decisions.¹⁷¹ Accordingly, BOE’s misuse of the Core Battery renders it invalid for in-service teachers as well.¹⁷² Notwithstanding that the Core Battery may have been properly validated for its original purpose—to screen teachers before they walk into a classroom—its misuse by BOE to adversely affect teachers who already were, and may still remain, in the classroom, renders the test invalid.

III. NO FURTHER EVIDENTIARY HEARINGS WOULD BE USEFUL OR PROPER

The trial record is damning for Defendant and this Court’s analysis requires no further evidentiary hearings. Even crediting the testimony of BOE and SED’s witnesses at trial, the record does not show the LAST “to be predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job for which Plaintiffs were being considered.”¹⁷³ No one responsible for the development of the LAST either followed proper professional methods in its validation or defined the harm the test was intended to prevent. In addition, there is no dispute that the LAST and Core Battery were misused. As a result, no credibility determinations are required by this Court. Moreover, to the extent BOE contends new testimony could so fundamentally re-shape the testimony at trial and alter Judge Motley’s Findings that formal validation could be established, this should be rejected out of hand. The trial in this matter concluded seven years ago and most events concerning validation

¹⁶⁹ Tr. 206-209; 261.

¹⁷⁰ Tr. 174; 977.

¹⁷¹ Tr. 1418-20.

¹⁷² See *Walston*, 566 F.2d at 1204.

¹⁷³ See *Gulino*, 460 F.3d at 385.

occurred twenty years ago. To suggest contrary or new evidence from original witnesses could now be extracted at a hearing is inherently incredible.¹⁷⁴ This proceeding can and should be decided on the existing record.

CONCLUSION

Plaintiffs respectfully request that this Court find (1) the LAST is not valid under the standards set forth in *Guardians*; (2) the LAST was misused when used for employment decisions regarding in-service teachers; (3) the Core Battery was misused when used for employment decisions regarding in-service teachers; and (4) that no further evidentiary hearings are required.

¹⁷⁴ See *Rosario v. Ercole*, 582 F. Supp. 2d 541, 601 (S.D.N.Y. 2008) (“Even where petitioner’s witnesses may have absolutely no intention of misleading the court, six-year old memories are inherently not as reliable as two and a half-year old ones.”); *Georgiadis v. First Boston Corp.*, 167 F.R.D. 24, 25 (S.D.N.Y. 1996) (“The passage of time always threatens difficulty as memories fade.”); *United States ex rel. Bisordi v. LaVallee*, 461 F.2d 1020, 1023-24 (2d Cir. 1972) (holding remand to district court in habeas petition would “serve no useful purpose” as the disputed lineup “occurred more than 12 years ago [and] [a]ll indications are that Ogilvie’s and Detective Carroll’s memory about the other men would be even dimmer today than at the trial in 1961.”); see also *Haouari v. U.S.*, 510 F.3d 350, 353 (2d Cir. 2007) (“It is axiomatic that witness recantations ‘must be looked upon with the utmost suspicion.’”) (citation omitted).

Dated: New York, New York
February 17, 2010

DLA PIPER LLP (US)

/s/ Joshua S. Sohn

Joshua S. Sohn
Anthony D. Gill
Spencer Stiefel
Rachel V. Stevens
1251 Avenue of the Americas
New York, New York 10020
(212) 335-4500

Anjana Samant
CENTER FOR CONSTITUTIONAL RIGHTS
666 Broadway, Seventh Floor
New York, New York 10012
(212) 614-6445

Stephen G. Seliger
Joel Hellman
155 North Michigan Avenue, Suite 501
Chicago, Illinois 60601
(312) 616-4244

Attorneys for Plaintiffs